Tellus
SERIES A
DYNAMIC
METEOROLOGY
AND OCEANOGRAPHY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

Tellus (2010), 62A, 30–47
Printed in Singapore. All rights reserved

TELLUS

Assimilation of image sequences in numerical models

By OLIVIER TITAUD[1,2]*, ARTHUR VIDARD[1,2], INNOCENT SOUOPGUI[2,1] and FRANÇOIS-XAVIER LE DIMET[2,1], [1]*INRIA Grenoble Rhône-Alpes, Montbonnot, 38334 Saint Ismier Cedex, France*; [2]*Université de Grenoble et CNRS, Laboratoire Jean Kuntzmann, BP 53, 38041 Grenoble Cédex 9, France*

(Manuscript received 28 October 2008; in final form 3 September 2009)

ABSTRACT

Understanding and forecasting the evolution of geophysical fluids is a major scientific and societal challenge. Forecasting algorithms should take into account all the available information on the considered dynamic system. The variational data assimilation (VDA) technique combines all these informations in an optimality system (O.S.) in a consistent way to reconstruct the model inputs. VDA is currently used by the major meteorological centres. During the last two decades about 30 satellites were launched to improve the knowledge of the atmosphere and of the oceans. They continuously provide a huge amount of data that are still underused by numerical forecast systems. In particular, the dynamic evolution of certain meteorological or oceanic features (such as eddies, fronts, etc.) that the human vision may easily detect is not optimally taken into account in realistic applications of VDA. Image Assimilation in VDA framework can be performed using 'pseudo-observation' techniques: they provide apparent velocity fields, which are assimilated as classical observations. These measurements are obtained by certain external procedures, which are decoupled with the considered dynamic system. In this paper, we suggest a more consistent approach, which directly incorporates image sequences into the O.S.

1. Introduction: images and assimilation

Understanding and forecasting the evolution of geophysical fluids (ocean, atmosphere, continental water) is a major scientific and societal challenge with important applications (prediction of extreme meteorological events, estimation of climate change, droughts and floods forecast, etc.). This problematic is currently the subject of an intensive research effort by the international scientific community. In order to achieve this goal, scientists need to take into account all the available information on the studied dynamic system. These informations can take several forms:

(i) *Numerical models:* They attempt to simulate the system dynamics on a wide range of time and space scales. Based on the Navier–Stokes equations, they are complicated because they must deal with the irregular shape of the domain, with the parametrization of unresolved processes, and with highly non-linear processes at the mesoscale and below. Numerical models represent mathematical and physical informations under the form of coupled sets of non-linear partial differential equations (PDE).

(ii) *Observations:* By essence, models simulate simplified real physical processes with a limited number of variables. Moreover, despite their regular improvement, certain model parameters are poorly known (e.g. initial and boundary conditions). This is why it is important to observe the system as well. Due to both satellite and in situ observing systems, observations have become increasingly numerous. Yet, their accuracy is not always satisfactory, and the processing of such a large quantity of heterogeneous data is difficult. Observations only provide a partial view of reality, localized in time and space. Therefore, some of them are indirectly related to model variables.

(iii) *Statistical informations:* On the one hand, the two sources of information described previously contain errors that can be characterized. Some physical processes are simplified or even neglected and errors are introduced through numerical processing (discretization, rounded errors, etc.). Most of the time there are known deficiencies of the model and therefore information about the structure of the needed correction is available. On the other hand, observation errors are mainly due to so-called representativity errors. Indeed, observation can represent phenomena that are not simulated by the model (subgrid processes).

*Corresponding author.
e-mail: Olivier.Titaud@imag.fr
DOI: 10.1111/j.1600-0870.2009.00416.x

Most of the time, the statistics of the representativity error can be estimated. All these informations are of importance and should be taken into account.

(iv) *Qualitative informations:* Forecasters use their knowledge on the system and additional materials such as satellite images to interpret the output of numerical prediction systems and provide corrected forecasts.

Taken separately models and observations do not permit for a deterministic reconstruction of real geophysical flows. Therefore, it is necessary to use these heterogeneous but complementary sources of information simultaneously through so-called data assimilation methods. These are inverse modelling techniques based on the mathematical theory of statistical estimation. Their aim is to combine observations and prior model estimates, taking into account their respective statistical accuracies, in such a way that the combined estimate is more accurate than either source of information taken individually. In the early 1980s Le Dimet and Talagrand (1986) proposed to use optimal control techniques to perform this task. This method, commonly referred to as variational data assimilation (VDA), has now been adopted by most of the main meteorological operational centres (European Centre for Medium Range Weather Forecast, Japan Meteorological Agency, Météo France, etc.). In VDA, the 'analysis problem' is defined by the minimization of a cost function that measures the statistically weighted squared differences between the observational information and their model counterpart. The cost function is minimized with respect to the control variables and this is done iteratively using a gradient-type descent method.

Models are based on fluid dynamics equations and on thermodynamic laws. After discretization in space (finite difference and spectral methods), the model is represented by a system of PDE with about 500 millions variables, and this number will keep increasing in the future. Data come from numerous sources (in situ, radiosondes, aircrafts, geostationary and polar orbiting satellites, etc.) and they are heterogeneous in density, quality and nature.

During the last two decades, many satellites have been launched to improve our knowledge of the atmosphere and of the oceans by observing the Earth. They provide photographic images of the Earth system among other data. Image-type data could be used in a classical data assimilation scheme, using the grey levels as observations of radiances emitted or reflected by the system. Yet connecting model variables with this kind of measurement is often complex, and consequently pixel levels are not assimilated in practice. However, satellites provide sequences of images that show the dynamic evolution of meteorological or oceanic 'objects' such as fronts, clouds, eddies, vortices, etc. The dynamics that the human vision can easily detect in this kind of image sequence clearly has a strong predictive potential. Unfortunately, this information is not optimally used in conjunction with numerical models for the moment. Furthermore, in other applications images are the only source of observation.

The purpose of this paper is to present a brief introduction to image sequences assimilation in a geophysical context. It also defines a methodology based on the theoretical background of VDA that combines the information coming from image sequences and the one coming from the numerical model, a priori knowledge and classical observations in a consistent way. Two classes of image assimilation techniques are currently considered:

(i) *Pseudo-observation:* An apparent velocity field is estimated from an image sequence using certain image processing techniques. This estimated field can then be used as indirect observations in a classical assimilation scheme. This has been done in a meteorological context in Schmetz et al. (1993).

(ii) *Direct image sequences assimilation (DISA):* Image sequences are assimilated through an appropriate observation operator directly into the optimality system (O.S.). This idea follows a general trend to avoid pre-processing of observations before its assimilation. A mathematical modelling of the image sequence is necessary to perform this task. Some recent work used the pixel basis (Papadakis and Mémin, 2008a; Corpetti et al., 2009) but we suggest that more sophisticated image representations should be used for geophysical applications.

Variational assimilation has also been used in a computer vision context (Papadakis and Mémin, 2008b). Since it deals with image evolution models this study cannot be truly considered as direct image assimilation in the sense we defined previously. Nevertheless, such developments are of interest and should be taken into account in a geophysical context.

Using image sequences together with numerical models may present several difficulties:

(i) The state variables of the numerical models (e.g. wind, temperature, pressure, humidity in atmosphere modelling and current velocities, temperature, surface elevation, salinity in ocean modelling) are not directly measured by satellites. What is observed is more or less linked to those variables by diagnostic equations, which will have to be included in the assimilation procedure.

(ii) The physical processes that are observed are not always taken into account in the model (e.g. local convection in atmospheric modelling, ocean colour in ocean modelling). This situation may happen for other kinds of observation as well but it is likely to be systematic in the case of image sequences.

(iii) Images are bi-dimensional informations whereas physical processes of geophysical fluids are three dimensional. From this point of view, a major difference between the ocean and the atmosphere comes from their radiative properties. In the first case images give informations about the ocean surface whereas in the second case images integrate all the radiative profile of the observed atmosphere column. Besides, certain meteorological

structures (e.g. some kind of clouds) are located in specific layers. Assimilating informations about the dynamic evolution of these structures needs additional informations about the altitude of the observed processes (using temperature profile measurements for instance). When the radiative properties of the ocean make the images give information about the ocean surface only (likely about a few centimetres), the corresponding observed processes may not be well depicted by the model.

(iv) As for human vision, we can presume that the pertinent informations coming from an image are mainly brought by its discontinuities or high gradients. Unfortunately, numerical models have a tendency to smooth these discontinuities out.

Finally, there is an important difference between meteorological and oceanographical images: in the first case, we can expect the model to more or less simulates the evolution of the observed features in an image sequence (e.g. fronts and vortex motion) but the altitude where such evolution takes place is not usually well known. In the second case, there is no ambiguity about the localization (i.e. at the surface) of the observed features because devices are not able to get information deeper than a few centimetres.

The paper is organized as follows. Basic principles of VDA techniques are described in Section 2. Section 3 briefly presents some mathematical considerations about images. Section 4 browses the different techniques currently available to assimilate—in a regular assimilation scheme—velocity fields derived from image sequences (pseudo-observations). A method for direct variational image sequences assimilation using a higher level of interpretation than the pixel level is described in Section 5: this forms the original part of this paper. Preliminary results are presented in Section 6.

2. Variational data assimilation

2.1. Principle of VDA

Let us briefly present the basic principle of VDA. The state of the system is described by a 'state variable' $\mathbf{x}$ depending on time and space. It belongs to the variable space state $\mathcal{X}$: we have $\mathbf{x}(t) \in \mathcal{X}$. It represents the variables of the model (e.g. velocity, temperature, elevation of the free surface, salinity, concentration of biological or chemical species, etc.). The evolution of $\mathbf{x}$ is governed by the differential system

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathcal{M}(\mathbf{x}, U), \quad t \in [0, T], \\ \mathbf{x}(0) = V, \end{cases} \tag{1}$$

where $U \in \mathcal{P}$ is an unknown set of parameters of the model: boundary conditions, model errors, parametrization of subgrid effects, etc. U may depend on space and time. The initial condition $V \in \mathcal{X}$ is unknown and depends on space. Let us assume that U and V being given, the problem (1) has a unique solution $\mathbf{x}(U, V, t), t \in [0, T]$. $\mathbf{y}(t) \in \mathcal{O}$ are given observations of the system available between 0 and T. The space $\mathcal{O}$ is called the space of observations. To make the description simpler, we suppose the continuity in time of the observations and that U only depends on space. The discrepancy between the observations and the state variable intervenes in the following 'cost function':

$$J(U, V) = \frac{1}{2}\int_0^T \| \mathbf{y}(t) - \mathcal{H}[\mathbf{x}(U, V, t)] \|^2_{\mathcal{O}} \, \mathrm{d}t + \frac{1}{2} \| U - U_0 \|^2_{\mathcal{P}} + \frac{1}{2} \| V - V_0 \|^2_{\mathcal{X}}, \tag{2}$$

where $\mathcal{H}$ is the so-called 'observation operator'. It is a mapping from the space of the state variable towards the space of observations $\mathcal{O}$ where the comparison is carried out (for more details see Section 2.2). The second and the third terms are regularization terms in Tikhonov's sense (Tikhonov, 1963). They also allow to introduce *a priori* informations. It is important to point out that J involves three different norms that define the topological structure of the functional spaces $\mathcal{X}$, $\mathcal{P}$ and $\mathcal{O}$; in practice, norms come from an inner product and can take into account the statistical information by introducing the error covariance matrices. For instance, a version discretized in space of the cost function (2) may involve

$$\| V - V_0 \|^2_{\mathcal{X}} := (V - V_0)^* \mathbf{B}^{-1} (V - V_0), \tag{3}$$

where $\mathbf{B}$ is the background error covariance matrix. The superscript * symbolizes the adjoint operation.

The problem of VDA can be considered as the determination of the *analysed state* (U^a, V^a) minimizing J, that is verifying

$$(U^a, V^a) = \operatorname{argmin} J(U, V). \tag{4}$$

A necessary condition for (U^a, V^a) to be a solution of (4) is that

$$\nabla J(U^a, V^a) = 0 \tag{5}$$

which becomes a sufficient condition if J is convex. As a first approximation, we have to solve a problem of unconstrained optimization despite the fact that, for instance, certain quantities are known to be non-negative (such as humidity, salinity or concentration). From a practical and numerical point of view, (U^a, V^a) is estimated by a descent-type algorithm, that is, by finding the limit of the sequence

$$\begin{pmatrix} U_{k+1} \\ V_{k+1} \end{pmatrix} = \begin{pmatrix} U_k \\ V_k \end{pmatrix} + \lambda_k D_k, \tag{6}$$

where λ_k is the step size realizing the minimum of J along the direction of descent D_k deduced from the gradient of J. For computing the gradient ∇J with respect to the control variables (U, V), we introduce the 'adjoint variable' $\mathbf{p}$ as the solution of the adjoint model

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} + \left[\dfrac{\partial \mathcal{M}}{\partial \mathbf{x}}\right]^* \cdot \mathbf{p} = \left[\dfrac{\partial \mathcal{H}}{\partial \mathbf{x}}\right]^* \cdot [\mathcal{H}(\mathbf{x}) - \mathbf{y}], \quad t \in [T, 0], \\ \mathbf{p}(T) = 0. \end{cases} \tag{7}$$

After a backward integration of the adjoint model, the gradient of J is given by

$$\nabla J = \begin{pmatrix} \nabla_U J \\ \nabla_V J \end{pmatrix} = \begin{bmatrix} -\int_0^T \left(\frac{\partial \mathcal{M}}{\partial U}\right)^* \cdot \mathbf{p} dt \\ -\mathbf{p}(0) \end{bmatrix}. \tag{8}$$

The derivation of the system (7) can be found in Le Dimet and Talagrand (1986). The model (1) plus the adjoint model (7) along with the cost function (2) form the O.S. The adjoint variable has the same dimension as the state variable of the direct model.

2.2. The observation space and operator

Measures are partial—in space and time—estimations of continuous physical quantities. Moreover, observations that are actually used in data assimilation schemes are in general a sampling of filtered measures. The space of observations $\mathcal{O}$ is a subset of $\mathbb{R}^p$ where p is the number of observations. Observations $\mathbf{y}$ and state variables $\mathbf{x}$ are linked through the 'observation operator'

$$\begin{aligned} &\mathcal{H}: \ \mathcal{X} \ \rightarrow \mathcal{O} \\ &\quad \mathbf{x}(t) \mapsto \mathbf{y}(t) = \mathcal{H}[\mathbf{x}(t)], \end{aligned} \tag{9}$$

which maps the state variable space $\mathcal{X}$ onto the observation space $\mathcal{O}$. The observation space is equipped with an inner product whose definition takes into account some uncertainties about the measurements and some representativity errors. We have

$$\|\mathcal{H}(\mathbf{x}) - \mathbf{y}\|^2_{\mathcal{O}} = (\mathcal{H}(\mathbf{x}) - \mathbf{y})^* \mathbf{R}^{-1} (\mathcal{H}(\mathbf{x}) - \mathbf{y}), \tag{10}$$

where $\mathbf{R}$ is the observation error covariance matrix.

We talk about direct observations when the observed quantities belong to the set of model state variables, and about indirect observations otherwise. In the first case, $\mathcal{H}$ may define a projection or an interpolation from $\mathcal{X}$ onto $\mathcal{O}$. In the second case, evaluating $\mathcal{H}(\mathbf{x})$ may be difficult from both computational and mathematical point of views because $\mathbf{y}$ and $\mathbf{x}$ may be connected through complex physical laws. The observation operator should be regular enough to ensure that the partial derivative $\frac{\partial \mathcal{H}}{\partial \mathbf{x}}$, involved in the adjoint model, (7) exists. In the case of direct observations, this property is generally satisfied whereas, in the case of indirect observations, this could be far less true. For instance, let us take the case of an ocean model with state variables velocity current, surface elevation, temperature and salinity. Direct observations could be in situ measures of temperature and salinity in certain specific locations of the studied domain. Indirect observations could be measures of surface salinity inverted from microwave radiances coming from remote sensing captors. In that case, $\mathcal{H}$ should represent the relation between these microwave radiances and surface salinity.

3. Interpretation levels and mathematical definitions of an image

Our goal is to define a methodology based on the theoretical background of VDA (see Section 2) in order to combine the informations coming from image sequences and the ones coming from the numerical model, a priori knowledge and classical observations in a consistent way. A first step is to identify a suitable level of interpretation of the image and a suitable mathematical framework to process it: this problematic is often called 'image modelling'.

3.1. Numerical raw images

The acquisition device converts the radiations coming from the observed scene into an electric signal, which is digitalized to form the 'raw image'. The device contains a sensor that is composed of an array of photo-sites where photons are accumulated. A numerical raw image is an array of pixels whose values represent the quantity of photons the corresponding photo-site received during the acquisition procedure. Raw images pixels can then be considered as radiance measurements. Black and white images pixels are described by one single channel called 'grey level' whereas coloured images pixels are described by three channels (Red, Green, Blue). The intensity of a pixel is usually described by a number between 0 (black) and 1 (white). This number is stored in a binary format with a string of digits (bits) whose length, the 'pixel depth', depends on the acquisition device and the digitization procedure. For instance, METEOSAT images have a 8 bits pixel depth (256 grey levels) as it is the case for common grey-scale images. Other remote sensing devices can produce 10 or 12 bits pixel depth images [e.g. Moderate-Resolution Imaging Spectroradiometer (MODIS)]. Note that computers usually store single precision floating point numbers with 32 bits (64 bits in double precision). If we consider that a pixel intensity is a measurement, then it is represented less precisely than single precision floating point numbers. Finally, the 'spatial resolution' of a satellite image refers to the quantity of Earth's surface a single pixel covers.

However, from a 'mathematical' point of view, an image is commonly considered as a real valued function of two continuous real variables. This allows the use of powerful mathematical tools developed by the image processing scientific community. A 'numerical image' is the discrete version of a mathematical one that is finally manipulated by computers.

3.2. Levels of interpretation

At the 'pixel level', an image of a fluid in motion is a digital instantaneous expression of the state of the fluid through the radiation that it reflects (visible light) or emits (infrared light). It is then a snapshot of an 'apparent state' which may involve other physical quantities than the model simulates, like tracers

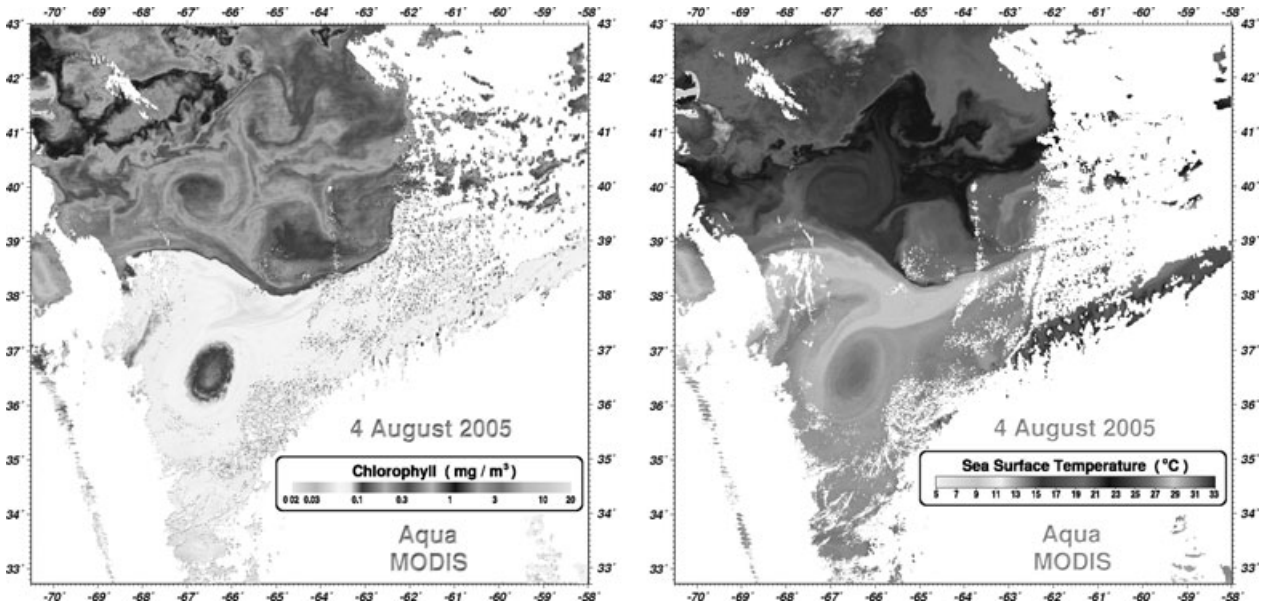

Fig. 1. MODIS-Aqua images collected on 4 August 2005 showing a couple of large Gulf Stream eddies in visible (left-hand panel) and infrared (right-hand panel) channels. In both image, we detect easily the same structures even though they are not linked to the same physical tracer radiance: chlorophyll concentration (left-hand panel) and surface temperature (right-hand panel). (Courtesy: NASA/ `http://oceancolor.gsfc.nasa.gov`).

(e.g. Cloud cover and aerosols in the atmosphere, chlorophyll in the ocean). The radiances are generally not part of the studied system and the dependence between this radiation and the state variables is not often trivial. For instance, 'cloud cover' is an intricate function of the state variable of meteorological models: clouds reflectance and intrinsic infrared emission depend both on thermodynamics (temperature and humidity) and on the microphysical processes that occur inside the clouds (involving water, ice, snow and the size of the particles). Satellite images that come from infrared sensors (see Fig. 1 right-hand panel) give indirect informations about the temperature of the observed system. At this level, information contained in images are not structured enough and represents a huge amount of data: typical satellite images are about 5000×5000 size, that is 25 millions of pixels. This number should be multiplied by the number of frames in the observed sequence. For this reason, this level will probably not be efficient in realistic image sequences assimilation. Note that low level vision problematics (denoising, deblurring, inpainting, segmentation, etc.) are usually performed at the pixel level.

At the 'analysis level', an image is symbolically described by its contents, typically by the objects and their characterization (geometry, layout, etc.) that appear in the observed scene. This kind of interpretation is used for high-level vision problematics as motion estimation, edge detection, pattern recognition, etc. The advantage of this level is that it may less depend on the effective radiance emission as suggests Fig. 1. In these images, the visible channel (left-hand panel) corresponds to a chlorophyll concentration map whereas the infrared channel (right-hand panel) corresponds to a surface temperature map. But in both of them, we may detect the same structures such as eddies and filaments. Currently, going from pixel level to analysis level in fluid flow image processing is a tough and still challenging task but this is beyond the topic of this paper. From now on we will use 'structure' as a generic word for the different features that may compose a geophysical image at the analysis level of interpretation.

3.3. Mathematical definition

There exist several mathematical definitions for images and it is not the aim of this section to describe them all. They essentially depend on the way the image needs to be processed. We can cite the multiscale decomposition approach, which comes from the signal processing theory [*wavelet* (Mallat, 1998) and the more recent *curvelet* (Candès and Donoho, 2004) transforms]. Another approach uses the PDE framework (Aubert and Kornprobst, 2006), where the image is considered as a function of two variables verifying a PDE equation that depends on the considered image processing problem. Certain applications use a stochastic approach as well to define images (Geman and Geman, 1984). For a good review of image analysis and the underlying mathematical definition, see Chan and Shen, 2005. In Section 6, we present an example where the anisotropic multiscale decomposition by curvelet transform is used.

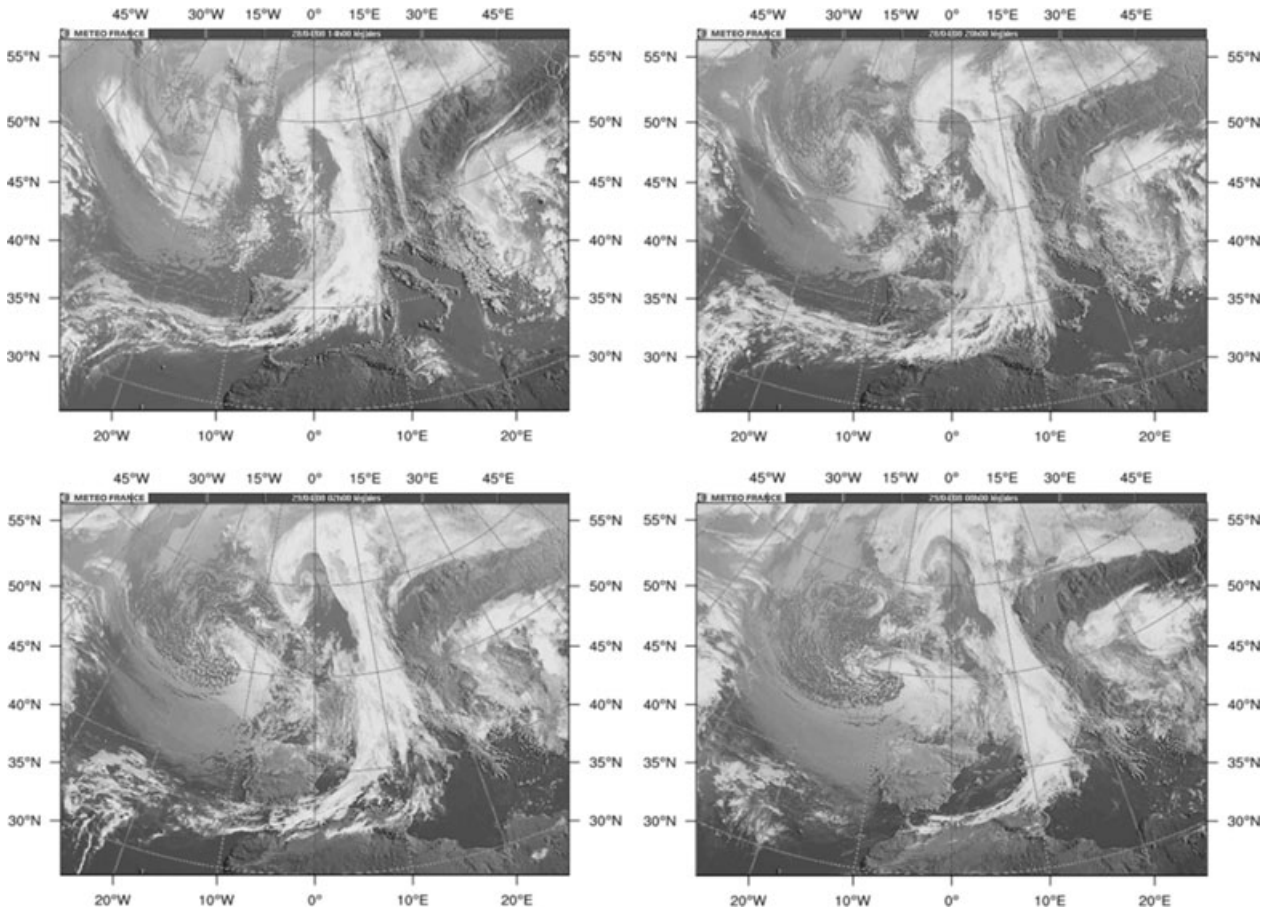

Fig. 2. Sequence of cloud coverage images above Europe from METEOSAT from 28 April 2008 to 29 April 2008. (Courtesy Météo-France).

In the VDA framework, the level of interpretation together with the mathematical definition of the image may be represented by the image-type observation space and its corresponding observation operator (see Section 5).

3.4. Sequence of images

With a static image, we only have access to Eulerian informations. Sequences of images provide dynamic informations, which can be interpreted as Lagrangian observations. The use of this type of information in VDA schemes was recently studied for ocean and hydrology problems (Nodet, 2005; Honnorat, 2007). Figure 2 is an image sequence taken over the Atlantic Ocean. In this example, we can see the evolution of a front over Western Europe associated with a cyclonic flow pattern over the ocean. This illustrates the ability of the human vision to estimate the evolution of the atmosphere through the features that it is able to detect and track. Because it takes into account all the observations in the assimilation window at once, the VDA framework seems well adapted to consistently process the dynamic information the image sequences provide. Processing and analysing image sequences may be done by extending the concepts listed above for static images using two different approaches: first, the sequence may be assimilated as a set of single two dimensional images (i.e. the frames) which are processed and analysed individually. Another approach is to consider the sequence as a three dimensional data: two spatial dimensions and time. The use of extensions of image processing techniques to the 3-D case may reinforce the temporal consistency of the DISA.

4. Images as source of pseudo-observation for data assimilation

Sequence of images are widely available in meteorology, oceanography, hydrology, astrophysics and even medicine. Historically, they have been mainly used for a 'by eye' analysis from experts in these fields. Yet, since the mid-1980s, research has been carried out to derive velocity fields from the sequences, with applications for fluid dynamics mainly (and very recently for movie compression and medical applications). We can use the velocity field derived from the image processing techniques

as pseudo-observation of the studied system, for instance in a regular VDA scheme (Herlin et al., 2004).

There are several ways to extract a velocity field from a sequence of images. They can be divided into two categories: the frame-to-frame motion estimators and the so-called image model technique (see Section 4.2).

4.1. Frame-to-frame motion estimator

The first class of motion estimator, directly based on image processing techniques, aims at estimating the velocity field that would 'transport' one frame to another.

Among them we can cite the statistical methods (mostly based on correlations between two images). They are mainly used for particle image velocities (PIV) and particle tracking velocities (PTV) experiments, that is, for lab experiments in fluid mechanics where particles are added to the fluid in order to follow the flow (Adrian, 1991). They are easy to implement, but they can be quite expensive when the number of particles increases. Applications of such methods are not really in the scope of what we are studying here.

Variational methods are more adapted to image sequences in geophysics. The most common one is usually called 'optical flow'. This classical approach in computer vision is based on the conservation of the global luminance between two images (Horn and Schunck, 1981): let $I : \Omega \times \mathbb{R} \to \mathbb{R}$ be the luminance of the pixel, the optical flow is the vector $\mathbf{v}(x,y)$ verifying

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\partial I}{\partial t} + \nabla I \cdot \mathbf{v} = 0. \tag{11}$$

This is not the only law of conservation, which could be considered according to the nature of the image: with an image of the colour of the ocean, an equation of conservation of chlorophyll (with source and sink terms) could be considered; with an image of sea surface temperature (SST), the Boussinesq approximation could be used.

Problem (11) is an ill-posed problem. Therefore, we need to add a regularization constraint. The choice of the regularization term depends on the application field [Vigan et al. (2000) for oceanography, Amodei and Benbourhim (1991) for wind field]. This problem can be summarized by finding $\mathbf{v}^\star$ that minimizes F defined by

$$F(\mathbf{v}) = \frac{1}{2}\int_\Omega \left\| \frac{\partial I}{\partial t} + \nabla I \cdot \mathbf{v} \right\|^2 \mathrm{d}x\ \mathrm{d}y + \frac{1}{2}\lambda \|S(\mathbf{v})\|^2, \tag{12}$$

where $S : \mathbb{R}^2 \to \mathbb{R}^2$ is a spatial regularization function (see Auroux and Fehrenbach, 2008) and $\lambda \in \mathbb{R}$ is the regularization factor. The minimization problem (12) is treated either by optimal control or by vector spline (Amodei, 1993; Suter, 1994; Isambert et al., 2007).

4.2. Image Model

Recently, Herlin et al. (2006), Huot et al. (2006) and Korotaev et al. (2007) proposed an extension of the optical flow methods called 'Image Model'. This is a more advanced motion estimator using techniques based on VDA which allows the use of several successive frames at once by adding a governing equation for $\mathbf{v}$ to the O.S. generating the pseudo-observations.

The aim is to retrieve a 'continuous' in time motion vector field from a sequence of images $I_{\mathrm{obs}}(t)$, $t \in [0, T]$. As in Section 4.1, it assumes the conservation of light and the image pixels are transported from one image to the other by the velocity field to be estimated. In order to improve the temporal consistency of the estimated velocities, this method requires a simple temporal evolution model $\mathcal{N}$. Along with the transport equation of the pixels (11), it forms the so-called Image Model:

$$\begin{cases} \dfrac{\partial I}{\partial t} + \nabla I \cdot \mathbf{v} = 0, \\ \dfrac{\partial \mathbf{v}}{\partial t} = \mathcal{N}(\mathbf{v}), \quad I(0) = I_0, \quad \mathbf{v}(0) = \mathbf{v}_0. \end{cases} \tag{13}$$

The pixel values of the images I_{obs} are then assimilated in the previous model using VDA (see Section 2.1) where the control variable is the initial condition $\mathbf{v}_0$ of $\mathcal{N}$.

$$F(\mathbf{v}_0) = \frac{1}{2}\int_0^T \|I - I_{\mathrm{obs}}\|^2 \mathrm{d}t + \frac{\lambda}{2}\|S(\mathbf{v}_0)\|^2. \tag{14}$$

The advantage of this method compared to the one described in Section 4.1 is that it takes into account all the images of the sequence at once and therefore can cope with missing data on some of the frames.

4.3. Limitations of the pseudo-observation technique

The velocity field produced by the motion estimation methods proposed in the two previous sections could be used as pseudo-observation in a classical VDA scheme such as presented in Section 2 (Herlin et al., 2004). However, despite their relative simplicity for implementation and their rapidity, pseudo-observation techniques based on motion estimation may suffer from some limitations. First, a frame to frame motion estimator does not take into account a sequence of more than two images. Due to the lack of consistency in time, it cannot capture the dynamic evolution of the image during the entire assimilation window. One of the consequences is that it cannot deal with missing data, which can be quite frequent in some of the applications targeted here. Missing data can be due to a glitch in the observing system or an obstacle (e.g. clouds).

The Image Model technique gives an answer to these two problems. Yet, the model $\mathcal{N}$ of the temporal evolution of the velocity only depends on $\mathbf{v}$ and therefore is likely to be a somewhat poor representation of the real behaviour of the velocity.

Neither of these two techniques takes into account physical informations about the underlying processes observed in the

images. Furthermore, some assumptions like the conservation of the luminance are clearly not valid for fluid flow images showing broken up structures. Finally, velocity fields obtained through the pseudo-observation approach are apparent velocities that can be quite different from the actual velocities. For instance, the absence of a tracer in a subset of an image sequence will lead to null apparent velocities in that area. Such techniques then need to provide error maps to account for this kind of error in the interpretation of the observation. Therefore, the definition of the observation operators and the specification of the corresponding observation error matrices (as for eq. 10) are likely to be a very difficult task. Finally, pseudo-observations that are taken into account in current operational centres are usually monitored by hand and the associated error is assumed to be large (Jean Pailleux, personal communication, 2008).

5. Direct image sequences assimilation

5.1. General framework

In the pseudo-observation approach, the information content is extracted from the images externally from the O.S. eqs (1), (2) and (7). Another way to proceed is a direct assimilation of the image sequence in the O.S. This approach is more consistent with the VDA philosophy, in particular the information content of the images is extracted consistently with the physics of the model. This technique intrinsically avoids the problem of the observation interpretation error mentioned above in the absence of tracers. In addition, for two decades, the trend is to avoid the inversion of the remote sensing data (i.e. radiances) before assimilation by including the direct radiative transfer model within the global assimilation scheme through observation operators (Talagrand, 1997). This leads to more suitable algorithms for error quantification.

The principle is to add a term directly linked to the images and their dynamics to the cost function. Let us denote by $\mathcal{M}_t(U, V)$ the state variable $\mathbf{x}(t)$ when the initial condition is $\mathbf{x}(0) = V$ and the model parameters are U. Then J would take the form:

$$J(U,V) = \underbrace{\int_0^T \|\mathbf{y} - \mathcal{H}[\mathcal{M}_t(U,V)]\|^2_{\mathcal{O}} \mathrm{d}t}_{\text{Classical term } J_o} + \int_0^T \|f(t) - \underbrace{\mathcal{H}_{\mathcal{X}\to\mathcal{F}}[\mathcal{M}_t(U,V)]}_{\substack{\text{Model to}\\ \text{Image}\\ \text{Operator}}}\|^2_{\mathcal{F}} \mathrm{d}t. \qquad (15)$$

In this expression $f(t) \in \mathcal{F}$ is a frame of an image sequence of the observed dynamic system at time t. It belongs to the space of images $\mathcal{F}$ equipped with a norm $\|\cdot\|_{\mathcal{F}}$. The operator $\mathcal{H}_{\mathcal{X}\to\mathcal{F}}$ is a mapping from the space of the state variable towards the space of images. It generates an image of the dynamic system at time t from the values of its state variable at this date: in the following $\mathcal{H}_{\mathcal{X}\to\mathcal{F}}$ is called 'synthetic image operator'. Note that for the sake of simplicity, we omit in (15) the regularization term which appears in the expression (2) of the classical VDA cost function.

The synthetic image operator works at the pixel level interpretation of the image (see Section 3). Papadakis and Mémin (2008a) showed that assimilating pixel levels directly gives a better result than using an optical flow-based pseudo-observation approach. Also in Corpetti et al. (2009) such formalism is used to assimilate images of pressure differences into a three-layer simplified atmospheric model. Nevertheless, we could point out two possible limitations of the use of pixel level for realistic applications in geophysics:

(i) A raw image sequence represents a huge amount of data. The size of a single satellite image is 5000×5000 pixels for MOP/MTP[1] visible channel. Consequently, a sequence of ten MOT/MTP images represents 2.5×10^8 data. This is therefore one or two order of magnitude greater than the 10^6– 10^7 current number of observations assimilated daily in operational NWP VDA system. Furthermore, the size of satellite images increases rapidly (for instance MSG-HRV[2] Channel images are currently about $6000 \times 12\,000$ pixels) and several channels may be taken into account.

(ii) Image resolution is often higher than the resolution of models. Visible channel MSG-HRV images have got a resolution of about 1 km (3 km for the infrared channels) whereas the horizontal resolution of the French ARPEGE global meteorological model varies between 15 and 90 km. MODIS satellite provides images with spatial resolution varying between 250 and 1000 m, depending on the considered spectral band. The global operational oceanographic system MERCATOR-PSY3V2[3] uses a model configuration with a horizontal resolution of $1/4^\circ$ (6–26 km). The use of the pixel level may then lead to some increase in the representativity error. Note that small-scale structures may be the consequence of mesoscale dynamic processes. This is particularly true for mixing tracer processes where spirals and filaments may come from the variability of mesoscale velocity field as it is clearly shown in Lehahn et al. (2007). In a pixel basis, those filaments are not represented as coherent structures while another mathematical modelling could avoid this problem.

Finally, the use of higher level of interpretation may also help to filter out certain unwanted signals (e.g. unrepresented processes or noise) from the data.

At the analysis level (see Section 3), an image is characterized by coherent structures (eddies, fronts, etc.). If we suppose that these structures can be approximated in a functional space $\mathcal{S}$ equipped with a suitable norm $\|\cdot\|_{\mathcal{S}}$, then the comparison

[1] Meteosat Operational/Transitional Program

[2] Meteosat Second Generation - High Resolution Visible

[3] `http://www.mercator-ocean.fr/html/mercator/index_en.html`

between image-type observation and model output can be done in that space and the cost function becomes:

$$J(U,V) = \int_0^T \|\mathbf{y} - \mathcal{H}[\mathcal{M}_t(U,V)]\|^2_{\mathcal{O}} \mathrm{d}t + \int_0^T \| \underbrace{\mathcal{H}_{\mathcal{F}\to\mathcal{S}}[f]}_{\substack{\text{Image to}\\ \text{Structure}\\ \text{Operator}}} - \underbrace{\mathcal{H}_{\mathcal{X}\to\mathcal{S}}[\mathcal{M}_t(U,V)]}_{\substack{\text{Structures}\\ \text{Observation}\\ \text{Operator}}} \|^2_{\mathcal{S}} \mathrm{d}t, \tag{16}$$

where $\mathcal{H}_{\mathcal{F}\to\mathcal{S}}$ 'extracts' structures from images (see next subsection): it represents the analysis level of the image. The operator $\mathcal{H}_{\mathcal{X}\to\mathcal{S}}$ observes the same type of structures from model state variables (model outputs). The norm $\|\cdot\|_{\mathcal{S}}$ depends on the mathematical definition of the image at the level $\mathcal{S}$. By choosing $\mathcal{S} = \mathcal{F}$ and $\mathcal{H}_{\mathcal{F}\to\mathcal{S}}[f] = f$, eq. (16) becomes equivalent to formulation (15).

With the usual variational approach described in Section 2, a term in the form

$$\left[\frac{\partial \mathcal{H}_{\mathcal{X}\to\mathcal{S}}}{\partial \mathbf{x}}\right]^* . (\mathcal{H}_{\mathcal{X}\to\mathcal{S}}[\mathbf{x}] - \mathcal{H}_{\mathcal{F}\to\mathcal{S}}[f]) \tag{17}$$

is added in the right-hand side of the adjoint model (7). In practical applications, we will then prefer working with differentiable structures observation operators. The differentiability of the observation operator depends on the observed quantity and this problem remains true for other kind of observation (not image) when they are too indirectly connected to model variables. Note finally that we may consider assimilating several image sequences (and possibly at different levels of interpretation) by adding corresponding terms in the DISA cost function (16).

5.2. Image to structure operator $\mathcal{H}_{\mathcal{F}\to\mathcal{S}}$

As we previously said, the image to structure operator represents the level of analysis at which we consider to manipulate images. It could represent transformations we apply to the image. The image space (in the mathematical sense) of this transformation is a space of another data type. The constraint which is imposed for choosing $\mathcal{S}$ is that it should have a structure of normed space to use simple rules for the differentiation of J. Moreover, its dimensionality must remain reasonable to store the extracted data.

Many different approaches for defining the space $\mathcal{S}$ exist. For instance, we can consider

(i) *Frequency characteristics:* in that case, $\mathcal{S}$ is the image space of a multiscale transformation such as wavelet and curvelet transforms. This approach is interesting for its well-known efficiency in data compression, denoising and edge extraction.

(ii) *Geometric characteristics:* certain image analysis techniques can detect certain geometrical features in the image (e.g. by means of active contours). This feature may take the form of a parametrized curve in $\mathbb{R}^2$ which could be assimilated as Lagrangian observation. This approach will necessitate high-level image analysis.

(iii) *Qualitative characteristics:* in geophysics, there are several recognizable structures connected to physical phenomena (e.g. cyclone, front, etc.). We could define $\mathcal{S}$ as a dictionary of such couples observable structure/physical phenomenon. More specifically in meteorology, clouds may give information about the flow or the physical processes where they are located. For instance, the particular shape of the stratus clouds is due to a stratified flow whereas cumulonimbus are known to involve strong convective processes.

5.3. Structures observation operator $\mathcal{H}_{\mathcal{X}\to\mathcal{S}}$

Features in geophysical images such as eddies or filaments may correspond to tracers in the observed system. A tracer can be defined as any quantity that is transported by the fluid flow. For instance, potential vorticity, temperature, salinity can be considered as atmospheric and ocean tracers. Modelling and simulating their evolution from the model state variables may help to define the structures observation operator. Indeed, we may use their description maps (e.g. concentration maps) as a 'synthetic image sequence' from which structures can be extracted using the image to structure operator. More precise correspondence between the description of the tracer and measured radiances (i.e. pixel intensity) may also be considered by modelling the acquisition procedure. For instance, this may properly connect the three dimensional information of the physical model with the two dimensional information of an image. Yet, as Fig. 1 suggests, we may think that such precise modelling is not always useful. Using a tracer evolution modelling leads to a structures observation operator of the form

$$\mathcal{H}_{\mathcal{X}\to\mathcal{S}} = \mathcal{H}_{\mathcal{F}\to\mathcal{S}} \circ \mathcal{H}_{\mathcal{X}\to\mathcal{F}}, \tag{18}$$

where $\mathcal{H}_{\mathcal{X}\to\mathcal{F}}$ is the synthetic image sequence operator. Note that if tracers are included in the model, the synthetic image operator is just some kind of projection (to be composed with an image acquisition model if necessary).

Yet certain features of interest in geophysical fluid images do not correspond to any tracers (such as waves or dry intrusions in atmosphere images). In that case, other techniques should be developed to extract structures from model outputs without synthetic image sequence production.

More specifically in the meteorological context, relevant structures may be located in specific layers of the atmosphere. Assimilating informations from the dynamic evolution of these structures requires complementary informations about the altitude of the observed processes. In the oceanographical context, filaments of phytoplankton that appear in ocean colour images are not only the result of the topological structure of the current field: they may also depend on vertical processes. Any pseudo-observation technique would require an a priori

knowledge of the altitude corresponding to the observed processes. Anyway it is unable to deal with patterns that cannot readily be attributed to a single vertical level. DISA is better suited to tackle this problem. Indeed the underlying 3-D physical processes are part of the O.S. through the model and the observation operator.

6. Numerical example

6.1. Experimental framework

Laboratory experimentation can be a good testbed for the new technique suggested in Section 5. Indeed, it offers the complexity of true physical processes while keeping it manageable. More importantly, it opens the way to realistic applications with images.

We use the same experimental framework as the study of the drift of a vortex on a turntable. The evolution of a vortex in the atmosphere is simulated at the CORIOLIS experimental turntable (Grenoble, France), which recreates the effect of the Coriolis force on a thin layer of water. A complete rotation of the tank takes 60 s, which corresponds to one Earth rotation. The vortex is created by stirring the water and it is made visible thanks to the addition of a passive tracer (fluorescein). Photographs of the vortex are taken from above the turntable and constitute the observed image sequence. For more details on the statements and the original motivations of this experiment, see Flór and Eames, 2002.

In this configuration, the evolution of the fluid can be modelled by the state variable $\mathbf{x} = (u, v, h)$ whose components verify the shallow-water equations

$$\begin{cases} \partial_t u - (f+\zeta)v + \partial_x B = -ru + \nu\Delta u, \\ \partial_t v + (f+\zeta)u + \partial_y B = -rv + \nu\Delta v, \\ \partial_t h + \partial_x(hu) + \partial_y(hv) = 0. \end{cases} \tag{19}$$

Unknowns are the zonal component $u(t, x, y)$ and meridional component $v(t, x, y)$ of the current velocity and the surface elevation $h(t, x, y)$. They depend on time t and the two horizontal directions x and y. We define the relative vorticity $\zeta = \partial_x v - \partial_y u$ and the Bernoulli's potential $B = gh + \frac{1}{2}(u^2 + v^2)$, where g is the gravity. The Coriolis parameter on the β-plane is given by $f = f_0 + \beta y$, ν is the diffusion coefficient and r the bottom friction coefficient. In this paper, the following numerical values are used for the parameters: $r = 0.9 \times 10^{-7}$ s^{-1}, $\nu = 0$ m^2 s^{-1}, $f_0 = 0.25$ s^{-1}, $g = 9.81$ m s^{-2} and $\beta = 0.0406$ m^{-1} . s^{-1}. The simulation is performed on a rectangular domain $\Omega =]0, L[\times]0, H[$ representing a subdomain of the turntable with $L = H = 2.525$ m. The domain is discretized on a $N \times N = 128 \times 128$ uniform Arakawa C-type square grid. A finite difference scheme is used for space discretization. Time integration is performed using a fourth order Runge–Kutta scheme. The time step is set to 0.01 s of the turntable experiment, which may correspond to 14.4 s in the atmosphere.

6.2. Assimilation procedure

In our test, we consider the problem of recovering the 'initial state' of the fluid $V(x, y) = \mathbf{x}_0(x, y) = (u, v, h)(0, x, y)$ which constitutes our control variable. The observations are images and we do not assimilate any measurements of any components of the state variable. We use an image to structure operator of the type 'frequency characteristics' by means of a multiscale and multi-orientation transformation of the images: the only source of information we assimilate is a subset of the decomposition coefficients (see below for more details). We consider assimilation windows of 30 s (3000 time steps) and 3.5 s (750 time steps) for twin experiments and real experiments, respectively. These window lengths are equivalent to 12 and 3 h in the atmosphere, respectively. Images are captured every 1.25 s (125 time steps) [resp. 0.25 s (25 time steps)]: this acquisition simulates a real capture frequency of one image per 30 min (resp. 6 min). Observations therefore consist of a set of 24 (resp. 30) images of 128×128 resolution, no observation being assimilated at initial time.

The application of VDA to the above test case using the formalism of Section 5 relies on the minimization of the following cost function:

$$J(\mathbf{x}_0) = \alpha_o \underbrace{\int_0^T \|\mathcal{H}_{\mathcal{F}\to\mathcal{S}}[f] - \mathcal{H}_{\mathcal{X}\to\mathcal{S}}[\mathcal{M}_t(\mathbf{x}_0)]\|_{\mathcal{S}}^2 \mathrm{d}t}_{J_o} + \alpha_b \underbrace{\|\mathbf{x}_0 - \mathbf{x}_b\|_{\mathcal{X}}^2}_{J_b}, \tag{20}$$

where observation and regularization terms J_o and J_b are balanced with two adjustable coefficients whose ratio α_b/α_o may represent the confidence in the background $\mathbf{x}_b$. The following subsections are devoted to a detailed description of these terms, for this experiment.

6.2.1. Observation term J_o.

6.2.1.1. Generation of synthetic images. As mentioned in Section 5.3, one way to define the structures observation operator $\mathcal{H}_{\mathcal{X}\to\mathcal{S}}$ is to first create an image from the model outputs using a synthetic image operator $\mathcal{H}_{\mathcal{X}\to\mathcal{F}}$ and then apply the image to structure operator $\mathcal{H}_{\mathcal{F}\to\mathcal{S}}$. In the current experimental framework, $\mathcal{H}_{\mathcal{X}\to\mathcal{F}}$ is obtained by simulating the advection of the monochromatic passive tracer into the fluid. To that purpose, we define

$$\mathcal{H}_{\mathcal{X}\to\mathcal{F}}[u, v, h] = q, \tag{21}$$

where $q = q(t, x, y)$ is the passive tracer concentration verifying the following conservation equation posed in Ω and for $t \geq 0$

$$\partial_t q + u\partial_x q + v\partial_y q = 0, \tag{22}$$

where (u, v) is given by (19). We obtain the synthetic image at a given date t using the concentration map $(x, y) \mapsto q(t, x, y)$ as grey levels.

6.2.1.2. Image to structure operator: curvelet transform and thresholding. As mentioned in Section 5.2, one way to extract relevant structures from the images is to work in the frequency domain by applying a multiscale transformation. In this paper, we use a curvelet transform (Candès and Donoho, 2004): an image f can be decomposed as

$$f = \sum_{j,k,l} \langle \varphi_{j,k,l}, f \rangle \varphi_{j,k,l}, \tag{23}$$

where $(\varphi_{j,k,l})_{j,k,l}$ are the elements of the curvelet frame. The inner product $\langle ., . \rangle$ is the usual $L^2(\mathbb{R}^2)$ one. The curvelet transform is a linear isometry and then its adjoint, needed in (7), is given by the inverse transformation (reconstruction). This is a convenient feature for practical use in the VDA framework. The use of curvelet transform is not the core of this method. Refer to Candès et al. (2006) and Ma and Plonka (2007) for more details on this subject and particularly for its discrete version.

In order to extract the structures from the images, we need to apply a threshold on the curvelet coefficients $\langle \varphi_{j,k,l}, f \rangle$. We therefore define the image to structure operator by the threshold of certain coefficients of the curvelet transform of the image f

$$\mathcal{H}_{\mathcal{F}\to\mathcal{S}}[f] = \mathcal{T}[\mathrm{DCT}(f)], \tag{24}$$

where $\mathcal{T}$ denotes a threshold operator. It is applied on the curvelet coefficients computed by the discrete curvelet transform (DCT) algorithm (taken from `www.curvelet.org`). In order to simplify the experiment, we use the following hard threshold function:

$$\tau(x) = \begin{cases} x & \text{if } |x| \geq \sigma, \\ 0 & \text{if } |x| < \sigma, \end{cases} \tag{25}$$

where x represents a coefficient in the curvelet series (23) and $\sigma > 0$ is a given threshold value. The space $\mathcal{S}$ is then the image space of the curvelet transform. We assimilate a fixed subset of the curvelet coefficients of the image decomposition. In our experiments, we keep only 10% of these coefficients. Using hard thresholding is more or less equivalent to assimilating images at the pixel level with the difference that it uses the efficient compression ability of this transformation (Candès and Donoho, 2004): this feature is critical for future realistic applications. A large choice of threshold functions is available depending on the application. Some classical threshold functions are listed in Antoniadis (2007). More sophisticated thresholding using a rule of total variation (TV) minimization can reduce the pseudo-Gibbs and element-like artefacts which appear with classical threshold functions (Durand and Froment, 2003). It may be used in this framework. DISA may give a starting point to very interesting investigations in the area of the application of curvelet and other multiresolution transformations (Ma et al., 2008).

6.2.1.3. Structures observation operator. The structures observation operator is the composition of the previous operator (24) with the synthetic image operator (21)

$$\mathcal{H}_{\mathcal{X}\to\mathcal{S}} = \mathcal{H}_{\mathcal{F}\to\mathcal{S}} \circ \mathcal{H}_{\mathcal{X}\to\mathcal{F}} = \mathcal{T} \circ \mathrm{DCT}. \tag{26}$$

6.2.2. Regularization term J_b. The regularization term

$$J_b(\mathbf{x}_0) = \|\mathbf{x}_0 - \mathbf{x}_b\|^2_{\mathcal{X}} \tag{27}$$

involves the background state $\mathbf{x}_b$ and a non-dimensionalization of the equations through the definition of the norm $\|\cdot\|_{\mathcal{X}}$. This norm is also used to define a pre-conditioning of the minimization problem.

6.2.2.1. Background term. In our experiments, the background is set to the constant value $\mathbf{x}_b = (u_b, v_b, h_b) = (0, 0, h_m)$ where h_m is the mean surface elevation in the tank. It approximately corresponds to a fluid at rest when the turntable speed is stationary. We then set the ratio α_b/α_o to a small value (about 10^{-6}). Note that in practical applications the first guess usually comes from a previous forecast of the system state. Initializing the assimilation process could also be possible with a velocity field deduced from an optical flow technique applied to the two first images of the sequence (initial surface elevation could be estimated from geostrophic balance). But we are not sure that these choices are relevant: results may indeed be too sensitive to them and what we want to point out is the feasibility of the method.

6.2.2.2. Preconditioning. The control state variable space $\mathcal{X}$ is equipped with the following norm:

$$\|\mathbf{x}\|_{\mathcal{X}} = \mathbf{x}^T \mathbf{B}^{-1} \mathbf{x}, \tag{28}$$

where $\mathbf{B}$ is the so-called background error covariance matrix. Its diagonal coefficients represent the standard error deviations of each variable with respect to itself: they are estimated from model outputs a priori. The velocity field is regularized using a method based on the Gradient Vector Flow technique (Souopgui et al., 2009). The off-diagonal blocks are the cross-covariances between variables. They come from balance relations between the observed variables and the others: this is a classical technique in VDA. Our observation operator connects images with the velocity field: this is why we have chosen to balance the surface elevation h with respect to (u, v) by 'inverting' the geostrophic equilibrium relation $u = -\frac{g}{f}\frac{\partial h}{\partial y}$, $v = \frac{g}{f}\frac{\partial h}{\partial x}$. Finally, we use the classical change of variable $\mathbf{x}'_0 = \mathbf{B}^{-1/2}\mathbf{x}_0$ for preconditioning. For more details about this strategy, see Vidard et al., 2004.

6.2.3. Optimization. The cost function J is successively minimized on assimilation subwindows of increasing sizes using the strategy described in Luong et al. (1998). This technique avoids the stagnation of the optimization algorithm around a local minimum. Such behaviour may otherwise occur because J is not convex. The minimization procedure is initialized

with the background (i.e. system at rest). In each successive minimization, we use the N1QN3 algorithm of the LIBOPT library (Gilbert and Jonsson, 2007).

6.3. Numerical results

As a first step to study the numerical feasibility of this method, we will focus on so called 'twin-experiments': observations do not come from the real world but are generated by the numerical model. The true state represents the evolution of a single isolated vortex subject to the Coriolis force. Assimilation of real experimental data coming from the CORIOLIS platform is shown in Section 6.3.2: the same configuration of the shallow-water model is used. Following the experimental set-up described above, the goal is to reconstruct the true initial state by assimilating only the image sequence and starting from a background at rest. From now on, dates and times are expressed in terms of their equivalent in the atmosphere.

6.3.1. Twin experiments. In this configuration, a 'true' state is generated by the shallow-water model (19) with the initial conditions of the experiment described in Flór and Eames (2002) (referenced AB): an isolated vortex is initialized with a radial velocity profile given by

$$v(r) = \frac{r v_m}{R_m} \exp\left[-\frac{1}{2}\left(\frac{r}{R_m}\right)^2\right], \tag{29}$$

in polar coordinate. R_m and v_m are the vortex radius and the maximum azimuthal velocity, respectively. In our experiments $R_m = 0.129$ m and $v_m = 0.02$ m s^{-1}. Initial surface elevation is generated using the geostrophic equilibrium equation. The mean surface elevation is set to $h_0 = 0.36$ m. True image sequence acquisition is simulated using synthetic images whose grey levels correspond to the concentration of a passive tracer which is advected by the velocity field (see Section 6.2.1). The boundary conditions imposed in the assimilation procedure slightly differ from the ones imposed during the synthetic observations generation: this simulates a model error.

Figure 3 (top panel) shows the superimposition of a zoom of the true velocity field and the corresponding true image at initial time (left-hand panel), at 12 h (centre panel) and at 24 h (right-hand panel). This period spans a 12 h assimilation and a

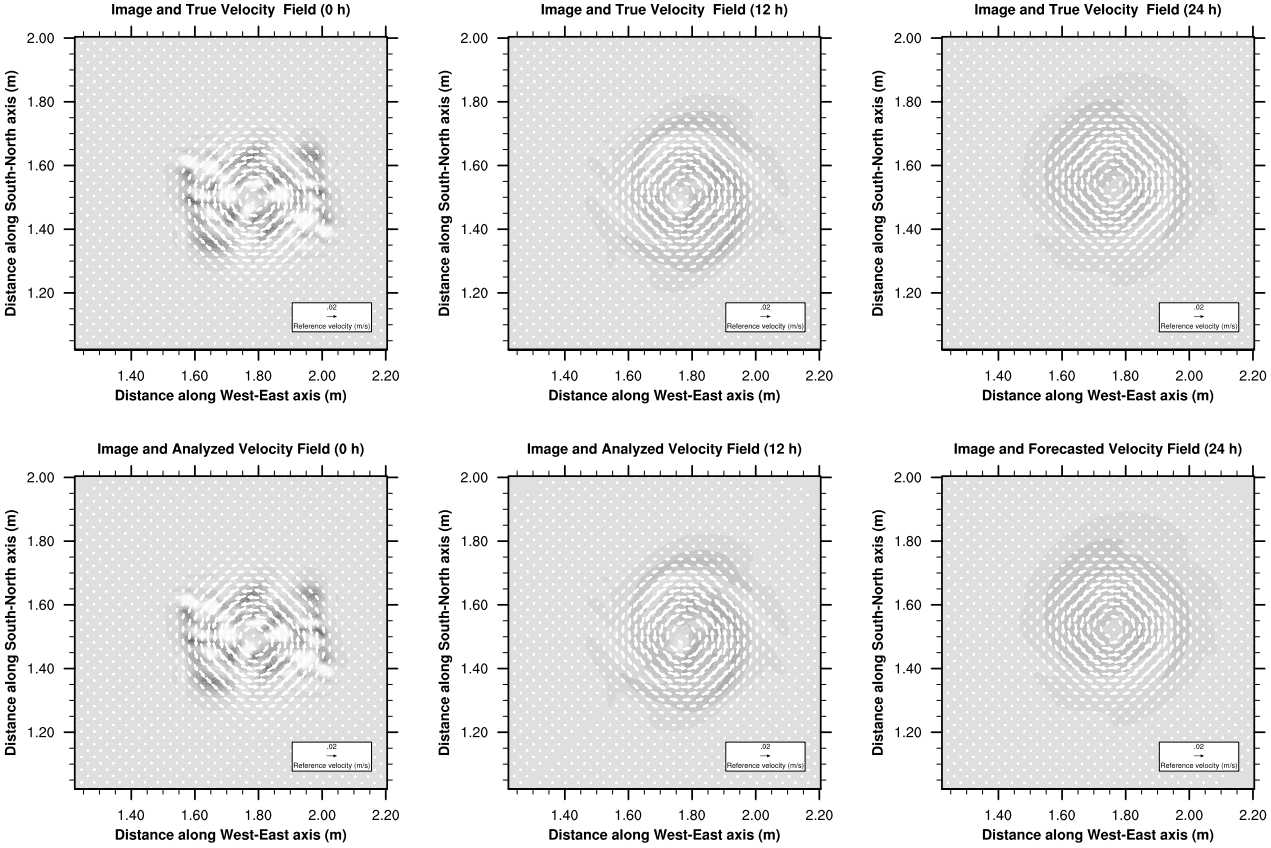

Fig. 3. Zoom of the true velocity fields over true images at initial time (0 h), at the end of the assimilation window (12 h) and at 24 h (top row). Analysed velocity fields over the corresponding true images at initial time (0 h) and at the end of the assimilation window (12 h). Zoom of the forecasted velocity field over the corresponding true image at 24 h (bottom row).

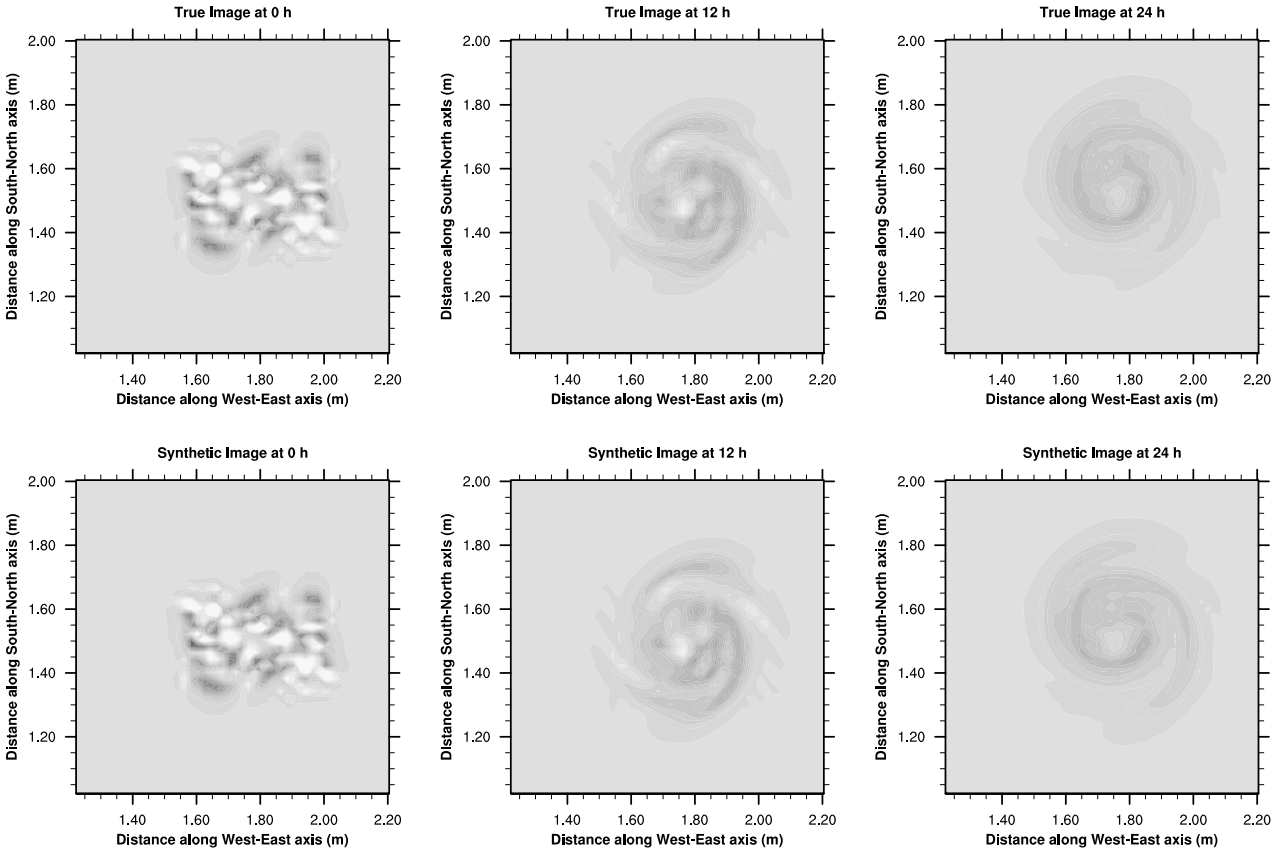

Fig. 4. Zoom of the true image sequence (top panel) and the corresponding synthetic image sequence (bottom panel) at 0, 12 and 24 h.

subsequent 12 h forecast windows. By comparison, Fig. 3 (bottom panel) shows the superimposition of the same zoom of the analysed velocity field and corresponding true images at initial time (left-hand panel) and at 12 h (centre panel). It also shows the superimposition of the same zoom of the forecasted velocity field and the corresponding true image at 24 h (right-hand panel). The background state is the system at rest (no current, flat surface). Figure 4 shows a zoom of the true image sequence at initial time (top left-hand panel), at 12 h (top centre panel) and at 24 h (top right-hand panel) and the corresponding synthetic image sequence reconstructed by the observation operator (bottom line). The differences are really small. The direct image sequence assimilation scheme therefore reconstructs a velocity field that properly transports the passive tracer. Consequently the assimilation scheme generates a synthetic image sequence that is close to the observed one. Figure 5 shows the evolution of the normalized root mean square errors of the velocity field $\mathrm{RMS}_{(u,v)}$ (left-hand panel) and of the surface elevation RMS_h

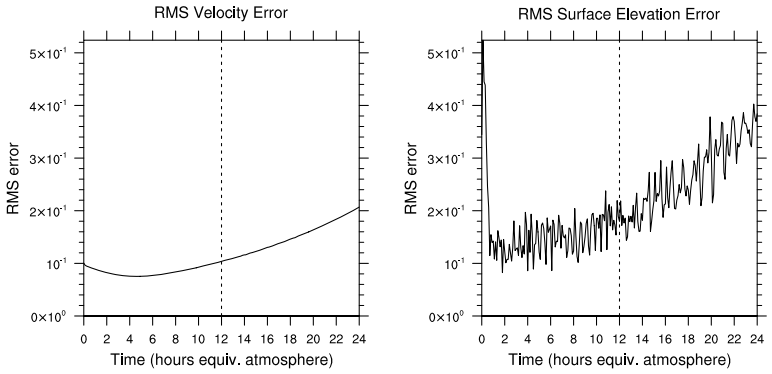

Fig. 5. Normalized root mean square errors for a 12 h assimilation and a subsequent 12 h forecast windows of the velocity field (left-hand panel) and of the surface elevation (right-hand panel). The dashed line represents the end of the assimilation window.

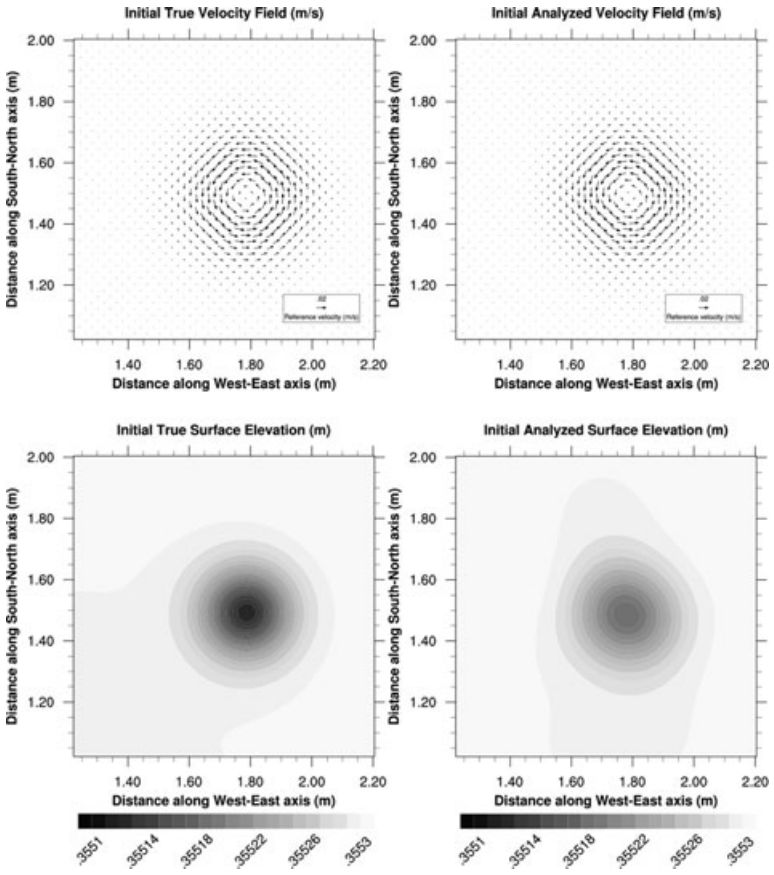

Fig. 6. Zoom of true and analysed state of the control variables: initial true velocity field (top left-hand panel), initial analysed velocity field (top right-hand panel), initial true surface elevation (bottom left-hand panel) and initial analysed surface elevation (bottom right-hand panel).

(right-hand panel) during the assimilation window (0–12 h) and during a subsequent forecast (12–24 h)

$$\mathrm{RMS}_{(u,v)} = \frac{\left[\sum_{i,j=1}^{N} \delta u^2(i,j) + \delta v^2(i,j)\right]^{1/2}}{\left[\sum_{i,j=1}^{N} u_t^2(i,j) + v_t^2(i,j)\right]^{1/2}}, \tag{30}$$

$$\mathrm{RMS}_h = \frac{\left[\sum_{i,j=1}^{N} \delta h^2(i,j)\right]^{1/2}}{\left[\sum_{i,j=1}^{N} (h_t(i,j) - h_0)^2\right]^{1/2}}, \tag{31}$$

where $\delta u = u_a - u_t$ is the difference between the analysed state u_a and the true state u_t. The dashed line delimits the end of the assimilation window. The behaviour of the Normalized RMS velocity error is expected: the U shape of the evolution of RMS error of the observed variables—indirectly in that case—is a diagnostic sign of the presence of model error. By comparison, the corresponding normalized RMS error of the surface elevation is large at initial time because the reconstructed surface elevation may not be well balanced. However, the model adjusts the surface elevation quickly. The surface elevation and the velocity field become better balanced and the RMS error dramatically decreases. This problem is a symptom of a weakness in the definition of the **B** matrix rather than specifically linked to the assimilation of image sequences. Figure 6 shows the true and analysed states of the control variables, that is, the initial velocity field and the initial surface elevation: the analysed states are the result of the minimization process applied to the cost function (20). Surface elevation seems to be underestimated but it is located at the right place. Comparison of velocity fields needs more quantitative tools. Figure 7 shows zooms of error maps between true and analysed initial velocity fields: the superimposition of the iso-contours of the Angular Error of Barron $AEB_{\mathbf{u}}$ between the true and analysed initial velocity fields and corresponding true velocity norm (left-hand panel) and the map of the relative error $\delta\|\mathbf{u}\|$ between the norms of the true and analysed initial velocity fields (right-hand panel). The Angular Error of Barron is defined by

$$AEB_{\mathbf{u}}(i,j) = \arccos\left(\left[\frac{\mathbf{u}_a(i,j)}{\|\mathbf{u}_a(i,j)\|}\right]^{*}\left[\frac{\mathbf{u}_t(i,j)}{\|\mathbf{u}_t(i,j)\|}\right]\right), \tag{32}$$

as it is suggested in Barron et al. (1994) and where $\mathbf{u} = (u, v)$ and $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^2$. The relative error $\delta\|\mathbf{u}\|$ is

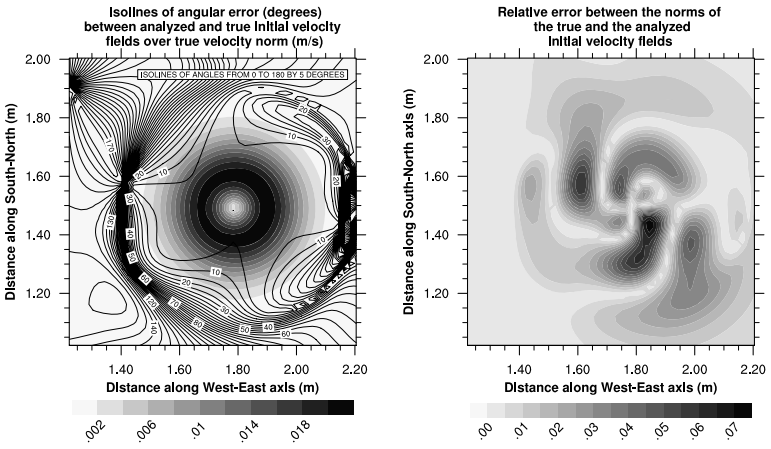

Fig. 7. Zoom of error maps between true and analysed initial velocity fields: iso-lines of angular error between true and analysed in initial velocity fields over the true velocity field norm (left-hand panel), relative error between the norm of true and analysed initial velocity fields (right-hand panel).

defined by

$$\delta\|\mathbf{u}\|(i,j) = \frac{\left|\|\mathbf{u}_a(i,j)\| - \|\mathbf{u}_t(i,j)\|\right|}{\max\limits_{i,j}\|\mathbf{u}_t(i,j)\|}. \tag{33}$$

Barron angular error is small in the area where the vortex is located. Error increases where information is missing that is at the boundary of the domain where the tracer is located (see Fig. 4). Note that angular error takes aberrant values where the velocity norm is insignificant and is not really relevant in those area. The relative error of the velocity field norm is quite small in general. It is not larger than 7%. The error map is structured because the background state is the system at rest. The structure shape may be explained by the irregular distribution of the tracer concentration.

To conclude, the direct variational assimilation of image sequence is able to reconstruct the initial vortex in the quite right position, magnitude and profile: this clearly shows the feasibility of this method.

6.3.2. Experiments with real images. We apply the same direct assimilation procedure as above to a real sequence of images showing the drift of a vortex on the CORIOLIS platform. This sequence comes from a similar experiment to the ones described in Flór and Eames (2002). In order to model this experiment we used the same shallow-water configuration as we used in the previous subsection. Yet few badly known physical parameters may be slightly different from the previous twin experiments. The rotation velocity of the vortex in this sequence is quite larger than the one we simulated previously. The image assimilation frequency is set to one image per 0.25 s to avoid aliasing phenomena. This frequency corresponds to the acquisition of one image per 6 m in the atmosphere. Figure 8 shows the superimposition of the analysed velocity field and the corresponding real image at initial time (left-hand panel), at 3 h which is the end of the assimilation window (centre panel) and the forecasted velocity field at 6 h. Figure 9 shows the analysed surface elevation at the same dates. The direct variational assimilation of this real experimental image sequence shows consistent results: it is able to reconstruct the anticyclonic secondary vortex at the southeast

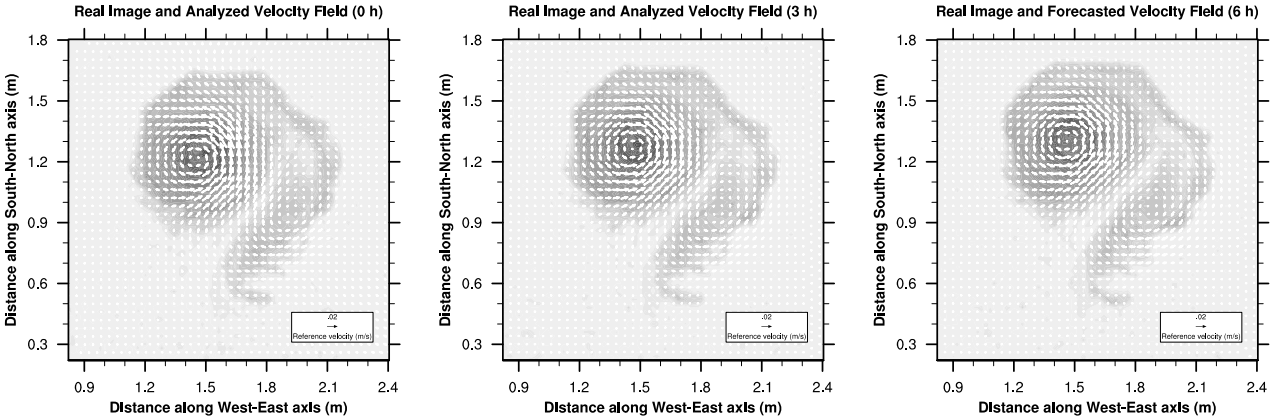

Fig. 8. Experiments with a real image sequence coming from the CORIOLIS platform. Superimposition of images and of a zooms of the analysed velocity field at initial time, at the end of the assimilation window (3 h) and the forecasted velocity field at 6 h.

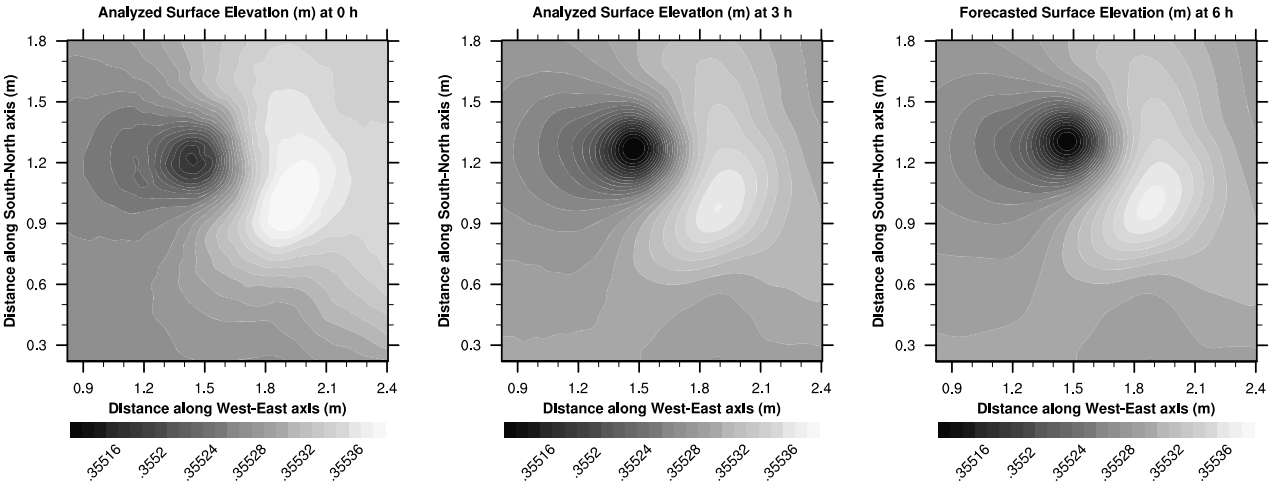

Fig. 9. Experiments with real images coming from the CORIOLIS platform. Zoom of the analysed surface elevation at initial time, at the end of the assimilation window (3 h) and the forecasted surface elevation at 6 h.

of the main one. The order of magnitude of both the velocity field and the surface elevation is also satisfactory and consistent with the experimental settings (J.B. Flór, personal communication, 2008). No other measures than images are available for this experiment, which prevent us from doing a more quantitative analysis of this result. This should be the goal of a further work with other experimental data.

7. Conclusions

Predicting the evolution of geophysical dynamic systems requires to link heterogeneous sources of information. There is a strong social demand for the improvement of the geophysical fluids prediction and VDA is a powerful tool to achieve this goal. The observation of the Earth by satellites is very expensive and, unfortunately, the dynamic information contained in image sequences is still underused by numerical forecasting systems. Image sequences represent a large amount of structured data. It is then of great interest to suggest some ways to process them in conjunction with numerical forecast systems.

Previous and present works could lead to the assimilation of data coming from image sequences into a VDA scheme. First, the so-called pseudo-observation technique assimilates velocity fields that are produced by image sequence analysis. Although efforts are made to improve the quality of these indirect observations, this technique intrinsically suffers from the lack of physical consistency with the observed system. Therefore, it is barely use in operational systems. Second, the DISA approach aims at processing the image sequences within the O.S. of the VDA scheme: it then combines image-type information consistently with the underlying physical model. Depending on the considered mathematical modelling, images may be processed at different levels of interpretation. The pixel level offers the natural way to exploit this kind of data by directly linking state variables to pixel values. A higher level of interpretation may avoid certain representativity error problem. A step further would be to work on both sides (model and observation) at a higher interpretation level, using model reduction techniques. For realistic applications in meteorology there is still a long way to go, alternatively we could start with projecting the modelled part of the observation operator only.

We presented a DISA method that is based on structure extraction from both image sequences and model outputs. We may use already existing processing and analysis image techniques to perform the structure extraction. Features in geophysical images usually correspond to the presence of tracers in the observed system. A natural way to extract structures from the model output is then to simulate the evolution of these tracers. This approach builds a synthetic image sequence from which the structures are extracted. Other techniques for extracting structures from model outputs avoiding synthetic image sequence production should be developed.

Our first results, both in a twin experiment and in a real data configuration, show that direct image sequences assimilation using structure extraction with anisotropic multiscale transformation of the images is achievable.

There is no doubt that investigating the way to optimally integrate image sequences in the VDA framework will bring up very interesting questions and will lead to improvements in both VDA and image analysis. We can expect that satellite image assimilation will improve the subgrid parametrization problematic in VDA since the resolution of a satellite image is often higher than the spatial resolution of numerical models. DISA may also initiate new definitions of image interpretations and improve specific feature tracking in geophysical fluid image sequences.

8. Acknowledgments

This work is part of the ADDISA project supported by the French National Research Agency (ANR) `http://addisa.gforge.inria.fr`.

References

Adrian, R. J. 1991. Particle imaging techniques for experimental fluid mechanics. *Ann. Rev. Fluid Mech.* **23**, 261–304.

Amodei, L. 1993. *A Study of a Vector Spline Function Class for the Approximation of a Velocity Field. Application to Meteorology.* PhD thesis, Université Paul Sabatier, Toulouse, France. (in French).

Amodei, L. and Benbourhim, M. N. 1991. A vector spline approximation. *J. Approx. Theory* **67**(1), 51–79.

Antoniadis, A. 2007. Wavelet methods in statistics: some recent developments and their applications. *Stat. Surveys* **1**, 16–55.

Aubert, G. and Kornprobst, P. 2006. Mathematical problems in image processing. In: *Mathematical Problems in Image Processing* Vol. 147, 2nd Edition (eds G. Aubert, and P. Kornprobst). Applied Mathematical Sciences Springer, New York.

Auroux, D. and Fehrenbach, J. 2008. Identification of velocity fields for geophysical fluids from a sequence of images, Research Report RR-6675, INRIA, Grenoble, France.

Barron, J., Fleet, D. and Beauchemin, S. 1994. Performance of optical flow techniques. *Int. J. Comput. Vision* **12**(1), 43–77.

Candès, E. and Donoho, D. 2004. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Comm. Pure Appl. Math.* **57**(2), 219–266.

Candès, E., Demanet, L., Donoho, D. and Ying, L. 2006. Fast discrete curvelet transforms. *Multiscale Model. Simul.* **5**(3), 861–899.

Chan, T. F. and Shen, J. 2005. Image processing and analysis. In: *Image Processing and Analysis* (eds T. F. Chan, and J. Shen). Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Corpetti, T., Héas, P., Mémin, E. and Papadakis, N. 2009. Pressure image assimilation for atmospheric motion estimation. *Tellus* **61A**, 160–178.

Durand, S. and Froment, J. 2003. Reconstruction of wavelet coefficients using total variation minimization. *SIAM J. Scient. Comput.* **24**(5), 1754–1767.

Flór, J.-B. and Eames, I. 2002. Dynamics of monopolar vortices on a topographic beta-plane. *J. Fluid Mech.* **456**, 353–376.

Geman, S. and Geman, D. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **6**, 721–741.

Gilbert, J. C. and Jonsson, X. 2007. LIBOPT: an environment for testing solvers on heterogeneous collections of problems—version 1.0, Technical Report RT-0331, INRIA.

Herlin, I., Le Dimet, F.-X., Huot, E. and Berroir, J.-P. 2004. Coupling models and data: which possibilities for remotely-sensed images, In: *e-Environment: Progress and Challenge of Research on Computing Science* Vol. 11 (eds P. Prastacos, U. Cortés, J.-L. D. D. León, and M. Murillo). Instituto Politécnico Nacional, Mexico, 365–383.

Herlin, I., Huot, E., Berroir, J.-P., Le Dimet, F.-X. and Korotaev, G. 2006. Estimation of a motien field on satellite images from a simplified ocean circulation model, *ICIP International Conference on Image Processing*, Atlanta, USA, 1077–1080.

Honnorat, M. 2007. *Lagrangian Data Assimilation for River Hydraulic Numerical Simulation*, PhD thesis, INP Grenoble. (in French).

Horn, B. and Schunck, B. 1981. Determining optical flow. *AI* **17**, 185–203.

Huot, E., Isambert, T., Herlin, I., Berroir, J.-P. and Korotaev, G. 2006. Data assimilation of satellite images within an oceanographic circulation model. *IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings.* **2**, 265–268.

Isambert, T., Berroir, J.-P. and Herlin, I. 2007. Fast and stable vector spline method for fluid apparent motion estimation. *IEEE International Conference on Image Processing, 2007. ICIP 2007.* **2**, 505–508.

Korotaev, G., Huot, E., Le Dimet, F.-X., Herlin, I., Stanichny, S. and co-authors. 2007. Analysis of the black sea surface currents retrieved from space imagery, *Rapport du 38ème Congrès de la Commission Internationale pour l'Exploration Scientifique de la mer Méditerranée*, Vol. 38, Istanbul, Turkey.

Le Dimet, F.-X. and Talagrand, O. 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus* **38A**, 97–110.

Lehahn, Y., d'Ovidio, F., Lévy, M. and Heifetz, E. 2007. Stirring of the northeast atlantic spring bloom: a Lagrangian analysis based on multisatellite data. *J. Geophys. Res.* **112**(C08005), doi:10.1029/2006JC003927.

Luong, B., Blum, J. and Verron, J. 1998. A variational method for the resolution of a data assimilation problem in oceanography. *Inverse Problems* **14**(4), 979–997.

Ma, J. and Plonka, G. 2007. Combined curvelet shrinkage and nonlinear anisotropic diffusion. *IEEE Trans. Image Process.* **16**(9), 2198–2206.

Ma, J., Titaud, O., Vidard, A. and Dimet, F.-X. L. 2008. Spatio-temporal structure extraction and denoising of geophysical fluid image sequences using 3D curvelet transforms, Research Report RR-6683, INRIA, Grenoble, France.

Mallat, S. 1998. A wavelet tour of signal processing. In: *A Wavelet Tour of Signal Processing* (ed. S. Mallat). Academic Press, San Diego, CA.

Nodet, M. 2005. *Mathematical Modelling and Lagrangian Data Assimilation for Oceanography*, PhD thesis, Université de Nice Sophia-Antipolis. (in French).

Papadakis, N. and Mémin, E. 2008a. Variational assimilation of fluid motion from image sequences. *SIAM J. Imaging Sci.* **1**(14), 343–363.

Papadakis, N. and Mémin, E. 2008b. A variational technique for time consistent tracking of curves and motion. *J. Math. Imaging Vision* **31**(1), 81–103.

Schmetz, J., Holmlund, K., Hoffman, J., Strauss, B., Mason, B. and co-authors. 1993. Operational cloud-motion winds from meteosat infrared images. *J. Appl. Meteor.* **32**.

Souopgui, I., Le Dimet, F.-X. and Vidard, A. 2009. Vector field regularization by generalized diffusion, Research Report RR-6844, INRIA, Grenoble, France.

Suter, D. 1994. Motion estimation and vector splines, In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 939–942.

Talagrand, O. 1997. Assimilation of observations: an introduction. In: *Data Assimilation in Meteorology and Oceanography: Theory and Practice* (eds M. Ghil, F. A. Bennett, P. Courtier, M. Kimoto and

N. Sato). WMO Second International Symposium on Assimilation of Observations in Meteorology and Oceanography. Meteorological Society of Japan and Universal Academy Press, Tokyo, 191–209.

Tikhonov, A. N. 1963. Regularization of ill-posed problems. *Dokl. Akad. Nauk SSSR* **4,** 1624–1627.

Vidard, A., Piacentini, A. and Le Dimet, F.-X. 2004. Variational data analysis with control of the forecast bias. *Tellus* **56A**, 177–188.

Vigan, X., Provost, C., Bleck, R. and Courtier, P. 2000. Sea surface velocity from sea surface temperature image sequences. 1. Method and validation using primitive equation model output.. *J. Geophys. Res.* **105**(C8), 19 499–19 514.